Origins of Protein Functions in Cells

Burckhard Seelig[1], Andrew Pohorille[2]*

[1]Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, St. Paul, MN 55108;
[2]Exobiology Branch, NASA Ames Research Center, Mail Stop 239-4, Moffett Field, CA 94035, USA

* andrew.pohorille@nasa.gov

In modern organisms proteins perform a majority of cellular functions, such as chemical catalysis, energy transduction and transport of material across cell walls. Although great strides have been made towards understanding protein evolution, a meaningful extrapolation from contemporary proteins to their earliest ancestors is virtually impossible. In an alternative approach, the origin of water-soluble proteins was probed through the synthesis and in vitro evolution of very large libraries of random amino acid sequences. In combination with computer modeling and simulations, these experiments allow us to address a number of fundamental questions about the origins of proteins. Can functionality emerge from random sequences of proteins? How did the initial repertoire of functional proteins diversify to facilitate new functions? Did this diversification proceed primarily through drawing novel functionalities from random sequences or through evolution of already existing proto-enzymes? Did protein evolution start from a pool of proteins defined by a 'frozen accident' and other collections of proteins could start a different evolutionary pathway? Although we do not have definitive answers to these questions yet, important clues have been uncovered.

In one example (Keefe and Szostak, 2001), novel ATP binding proteins were identified that appear to be unrelated in both sequence and structure to any known ATP binding proteins. One of these proteins was subsequently redesigned computationally to bind GTP through introducing several mutations that introduce targeted structural changes to the protein, improve its binding to guanine and prevent water from accessing the active center. This study facilitates further investigations of individual evolutionary steps that lead to a change of function in primordial proteins. In a second study (Seelig and Szostak, 2007), novel enzymes were generated that can join two pieces of RNA in a reaction for which no natural enzymes are known. Recently it was found that, as in the previous case, the proteins have a structure unknown among modern enzymes. In this case, in vitro evolution started from a small, non-enzymatic protein. A similar selection process initiated from a library of random polypeptides is in progress.

These results not only allow for estimating the occurrence of function in random protein assemblies but also provide evidence for the possibility of alternative protein worlds. Extant proteins might simply represent a 'frozen accident' in the world of possible proteins. Alternative collections of proteins, even with similar functions, could originate alternative evolutionary paths.

References

Keefe AD and Szostak JW (2001) Functional proteins from a random-sequence library, *Nature*, 410: 715.

Seelig B and Szostak JW (2007) Selection and evolution of enzymes from a partially randomized non-catalytic scaffold, *Nature*, 448: 828.